Distribution of Age at the First Marriage in Java-Bali

Hadi Sumarno

Abstract *Woman's age of marriage is one of the important proximate determinant of fertility. A number of research have been done to explain marriage patterns across age, cohort, and socioeconomic factors. One of the powerful models that often used to explain the marriage pattern is Coale-McNeil model. The purpose of this study is to determine the Java and Bali marriage patterns using Coale-McNeil marriage model. Data used are based on the 1991 Indonesia Demographic and Health Survey. The study shows that there is quite a great variation in mean age of first marriage across provinces. It was also found that the mean age of first marriage changes across cohorts. The younger cohorts tend to marry at older age.*

Keywords: Coale-McNeil marriage models; dispersion parameter; incidence of marriage; age at marriage.

1. Introduction

We realize that over population can increase both environmental and social problems for many countries. There are a lot of research in demography which have been done to explain the factors of the population growth. When the net migration usually is insignificant and mortality rates continue to fall recently, fertility plays an important role in determining the population growth. One of the important factors determining fertility is age at first marriage. Hajnal (1953) showed that the baby boom happens after marriage boom.

At least there are five measurement techniques that have been proposed to explain the marriage patterns, i.e., (1) proportion ever married,

(2) quartile and median age at marriage, (3) interquartile, (4) singulate mean age at marriage, and (5) Coale-McNeil marriage model. The last two techniques are commonly used in demographic research. The singulate mean age at marriage (SMAM) proposed by Hajnal in 1953. He employed life table approach in his technique. Coale-McNeil marriage model proposed by Coale and McNeil in 1972. This technique has some advantages compared to others. Among them are its abilities to estimate the incidence of marriage and the dispersion of marriage process and its parametric in nature. The latter makes it able to deal with a wide range of data; such as grouped or ungrouped data, household or individual survey based data, all women or ever married women sample only, and censored or uncensored data.

The purpose of this study is twofold. First, is to compare the different of marriage patterns across provinces. Second, is to see the trend of marriage pattern across cohorts.

2. Model

Coale-McNeil model is an eclectic model. It is a combination between theoretical, empirical, and conceptual model. From the empirical evidence of some populations, Coale (1971) found out that age at first marriage of any population has a similar pattern. They varied only in location (a_0), scale (k), and area under the curve (c). The location is interpreted as age of initially marriage. The scale indicates the speed of marriage processes in any population relative to the standard patterns. The area under the curve is the proportion of the cohort eventually marrying.

Theoretically, marriage can be seen as a convolution of the marriage processes, the age of entry into the marriage market and the delays involved in finding a suitable partner, being engaged, and becoming married. Fenney in Coale and McNeil (1972) suggested the conceptual models of these processes; the first process is normal and the next three processes are distributed exponentially. In order that the theory as explained earlier can be implemented mathematically, Coale and McNeil (1972) added one more assumption that the exponential delays have means in harmonic series.

Taking advantage of that information, Coale and McNeil (1972) discovered the distribution of age at first marriage of Sweden 1865-1869 as follows:

$$g_w(W) = 0.196\ exp\{[-0.174(W-6.06)] - exp[-0.288(W-6.06)]\}$$

with W is the mean of corrected random variable of Sweden's age at the first marriage. Suppose that Y is the random variable of age at the first marriage of any population. If f(y) is the proportion of married at age y, then f(y) can be explained as the multiplication between incidence of marriage (c) and probability of married at age y as

$$f(y) = c\ g(y)$$

$$g(Y) = \frac{1}{k} g_w\left(\frac{Y - y_0}{k}\right)$$

k = scale parameter
y_0 = initially marriage.

Rodriguez and Trussell (1980) said that these parameters were not easy to interpret. Then, they transformed the original standard schedule to another schedule using standardization transformation. The new standard schedule has mean zero and variance unity. The new standard schedule is

$$g_z(Z) = 1.2813 exp\ \{-1.145(Z+0.805) - exp\ [-1.896(Z+0.805)]\}$$

with $Z = (W - \mu_W)/\ \sigma_W$, μ_W and σ_W are mean and standard deviation of Sweden marriage pattern respectively. Then, the distribution of age at the first marriage of any population, Y, can be explained as

$$g(Y) = \frac{1}{\sigma_y} g_z\left(\frac{Y - \mu_y}{\sigma_y}\right)$$

with μ_y and σ_y are mean and standard deviation of random variable Y respectively.

3. Parameter Estimation

The estimation of the parameters can be done using maximum likelihood estimation. Technically, it depends on the nature of the data. Mainly there are two kinds of data set in demographic survey. First, data that only contain marital status for given age. This kind of data are usually also called as household sample survey data. Using this kind of data, we can work with cross sectional analysis only. Second, data that contain information of age of first married. This kind of data are usually also called as individual survey data. Using this kind of data, we can analysis the data based on cohorts.

In this study, we use the grouping data for both kinds of data set. The first kind of data is used to see the marriage pattern across provinces in Java-Bali. And the second kind of data is used to see the trend of marriage pattern across cohorts. When the data are grouped, the number of ever married women between age a_i and $(a_i+\Delta a)$, i=1,2, ..., k, are distributed binomially. The loglikelihood function (excluding the constant factors) become

$$L \propto \sum_{i=1}^{k} \left[m_i \ln(p) + s_i \ln(1-p) \right]$$

m_i and s_i are the number of ever married and never married women at age group i respectively. The loglikelihood function L is maximized with respect to p, with p = c G,

$$G = \int_{-\infty}^{a^*} \frac{1.2813}{\sigma_Y} \exp\!-\frac{1.145}{\sigma_Y}\left(y - \mu_Y + 0.805\sigma_Y\right) - \exp\left[-\frac{1.896}{\sigma_Y}\left(y - \mu_Y + 0.805\sigma_Y\right)\right] dy$$

a^* is the center point of the grouping criteria. A program in FORTRAN was generated by using E04JBF NAG routine to calculate μ_Y, σ_Y, and c. For another kind of grouped data (or even for individual data), it still can use this algorithm with appropriate distribution function on the likelihood function. Using the maximum likelihood function, the standard error of the parameters being estimated also can be obtained using information matrix (*see* Rodriguez and Trussell 1980).

To see the goodness of fit, the likelihood ratio test was used as follows:

$$LRT = -2\sum_{i=1}^{k}\left[m_i \ln\left(\frac{\hat{p}_i}{p_i}\right) + S_i \ln\left[\frac{1-\hat{p}_i}{1-p_i}\right]\right]$$

with $\hat{p}_i = \hat{F}(\hat{a}_i) = \hat{c}\hat{G}(\hat{a}_i)$ and $p_i = m_i/(m_i + s_i)$. For the large sample, LRT distributed chi square with *k-1* degree of freedom. In general, when the value of LRT $< \chi^2_{\alpha}$, with α is degree of significance, we can say that the model is able to fit the data well.

Unfortunately, using the likelihood ratio test, the statistic LRT increases are not only caused by deviation of the model from the data, but also due to the increasing number of observation. Although the deviation is small, when the number of observation is too large, the model will be rejected. Since this test usually rejects the model even it is good when the data is too large, so the Bayesian Information Criteria (BIC) was also employed as the alternative test as follow:

$$BIC = LRT - df\ ln(n)$$

with *df* is degree of freedom and *n* is the number of observation. When the Bayesian statistic is less than zero, the model is accepted (*see* Raftery, 1986; 1987, Xie, 1990, Xie and Pimentel, 1992).

4. Data

The data used were based on the 1991 Indonesia Demographic and Health Survey (1991 IDHS). The data were collected by the National Family Planning Coordinating Board, Central Bureau of Statistics, Ministry of Health, and Macro International. Using non proportional sampling, 26858 households were interviewed. There were 23470 eligible women and 22909 of them successfully interviewed. More detail information about this survey can be seen in CBS *et al.* (1992).

Sampling in Java and Bali was held in two stages. The first stage was choosing the sampling area proportional with urban and rural areas. After that, from each sampling area 25 households were randomly selected. The number of sample in Java and Bali contained 10617 households of 45615 persons. There were 8622 eligible women that were successfully interviewed.

5. Results

The result of analysis using the household data of Java-Bali is presented in Table 1. According to the likelihood ratio test, the Java-Bali data as a single population do not close to the model. It does not always mean that the marriage patterns in Java-Bali do not follow the Coale-McNeil marriage model. This result can happen due to the following reasons. The data are too large. As mentioned by Raftery (1986) the model often rejected when the number of data is large, although the model is good. Another reason is that the marriage patterns in Java and Bali do not have the same characteristics. Some women who live in the specific region may tend to marry in younger age for some reasons, because of cultural or economic conditions. Otherwise their counterparts who live in another region tend to marry in older age. It is supported by the result of analysis using separate data of each province as presented in the same table. It can be seen that the probability of likelihood ratio test in six provinces in Java-Bali is greater than 5 percent. It means that the women who live in the same provinces tend to have similar characteristics. Using the urban and rural data separately can increase the power of test.

Table 1
PARAMETER ESTIMATES
USING HOUSEHOLD SAMPLE DATA OF JAVA-BALI

Region	μ	σ	κ	SE μ	SE σ	SE κ	LRT	Pr. LRT	BIC
Urban+Rural									
Java-Bali	20.76	4.50	0.987	0.0864	0.0165	0.0137	54.19	0.009	-247.22
DKI Jakarta	24.25	6.11	0.993	0.2297	0.2604	0.0065	34.62	0.344	-221.86
West Java	19.32	3.43	0.989	0.1598	0.2042	0.0039	42.66	0.099	-203.31
Central Java	21.25	4.27	0.992	0.2136	0.2442	0.0045	27.31	0.703	-214.78
DI Yogyakarta	23.98	5.12	0.973	0.2937	0.0731	0.0398	26.82	0.726	-210.66
East Java	20.44	4.56	0.989	0.2042	0.2586	0.0045	25.75	0.775	-217.14
Bali	22.41	4.17	0.911	0.3099	0.0974	0.0415	44.75	0.067	-190.01
Urban									
Java-Bali	23.14	5.09	0.978	0.1220	0.0324	0.0174	45.08	0.062	-241.01
DKI Jakarta	24.25	6.11	0.993	0.2300	0.2609	0.0065	34.62	0.344	-221.86
West Java	21.45	3.70	0.974	0.2916	0.0759	0.0407	37.74	0.223	-179.40
Central Java	24.09	5.88	0.997	0.5726	0.6678	0.0166	27.16	0.710	-175.62
DI Yogyakarta	24.87	5.09	0.950	0.1844	0.4756	0.0457	25.80	0.772	-187.58
East Java	23.37	5.08	0.977	0.4355	0.1096	0.0597	17.56	0.982	-189.90
Bali	23.84	4.74	0.911	0.9846	0.2411	0.0723	33.44	0.397	-158.99
Rural									
Java-Bali	19.26	3.42	0.993	0.0668	0.0826	0.0013	90.58	0.000	-210.85
West Java	18.20	2.73	0.998	0.1717	0.2184	0.0023	37.53	0.230	-191.61
Central Java	20.27	3.49	0.996	0.2074	0.2213	0.0030	20.19	0.948	-219.72
DI Yogyakarta	22.82	4.69	0.983	0.3737	0.0721	0.0511	25.64	0.779	-191.42
East Java	19.15	3.56	0.995	0.2012	0.2468	0.0030	30.47	0.544	-199.58
Bali	21.87	3.88	0.911	0.3440	0.0993	0.0459	44.43	0.071	-180.30

Notes: SE = Standard Error
LRT = Likelihood Ratio Test
Pr.LRT = Probability test of LRT
BIC = Bayesian Information Criteria

The BIC test also supports this result. From the last column it can be seen that all of the BIC statistics are less then zero, even for the overall Java and Bali data. The different result between the LRT test and BIC test has been explained above. The BIC test is the corrected LRT test, to reduce the bias of the number of observations which is too large. As can be seen in the first row of Table 1, the BIC test accepts the model that is rejected by LRT test.

The result of analysis using data of all women sample presented in Table 2. All the probabilities of likelihood ratio test from this analysis are equal zero, means that the LRT test does not suitable to be used, because of the sample is too large. According to this reason we only present the BIC values. From the Table 2 it can be seen that all of the BIC values are less

than zero. It means that the women who live in the same cohort tend to have the similar pattern of the age at the first marriage.

Table 2
PARAMETER ESTIMATE USING THE ALL WOMEN SAMPLE DATA OF JAVA-BALI

Region	μ	σ	κ	SE μ	SE σ	SE κ	BIC
Urban							
Birth Cohort							
1962-1966	20.99	5.89	0.895	0.3011	0.5063	0.0210	-218.7
1957-1961	19.36	5.00	0.944	0.1981	0.3372	0.0101	-267.0
1952-1956	19.33	5.03	0.967	0.2071	0.3500	0.0079	-408.9
1947-1951	18.04	4.41	0.957	0.2169	0.3603	0.0099	-562.8
1942-1946	17.63	4.30	0.975	0.2033	0.0612	0.0415	-866.4
Rural							
Birth Cohort							
1962-1966	17.98	4.37	0.997	0.1683	0.2850	0.0080	- 75.4
1957-1961	17.06	4.05	0.994	0.1522	0.1292	0.0044	-250.8
1952-1956	16.68	3.83	0.991	0.1520	0.1266	0.0012	-520.8
1947-1951	16.14	3.52	0.991	0.1536	0.0334	0.0158	-719.5
1942-1946	16.27	3.63	0.988	0.1540	0.1271	0.0045	-711.8

Notes: SE = Standard Error
BIC = Bayesian Information Criteria

6. Discussion

Table 1 shows that almost all women in Java-Bali ever married at least once during their lifetime. The mean of age at marriage is 20.76 years. Those who lived in rural areas had age of first marriage four years younger compared to women who lived in urban areas. Women who lived in rural areas also tended to marry in the shorter range of age, relative to their counterpart who lived in urban areas (σ_{rural} =3.42 years, σ_{urban} =5.09 years). It means that when her friends are going to marry, she also tends to marry soon. She is afraid to be called an old-unmarried girl. More clearly, the distribution of marriage patterns of women in Java-Bali also can be seen in Figure 1.

Figure 1
DISTRIBUTION OF WOMEN'S AGE AT THE FIRST MARRIAGE IN JAVA-BALI

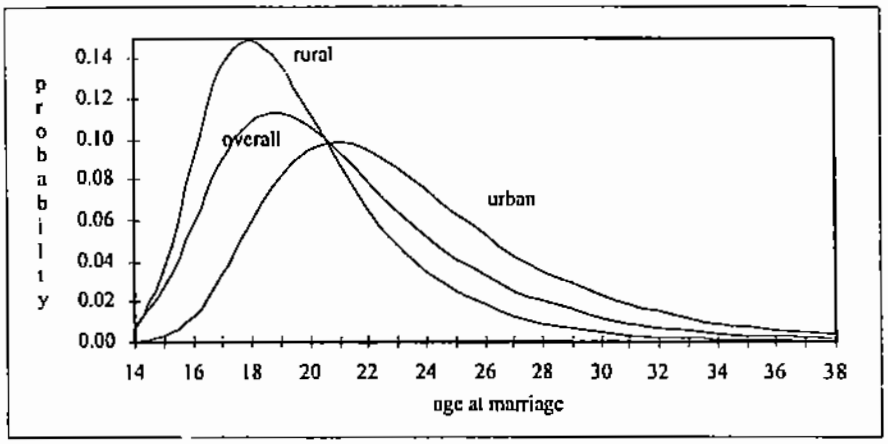


Table 1 also shows that there are high variations in age at first marriage across provinces, from 19.32 years for West Java, followed by East Java 20.44 years, Central Java 21.25 years, Bali 22.41 years, DI Yogyakarta 23.98 years, and DKI Jakarta 24.25 years. DKI Jakarta has the highest mean has age of first marriage because all of the women live in urban areas. The similar reason is for DI Yogyakarta. All of the rural women in DI Yogyakarta live next to cities around DI Yogyakarta including Yogyakarta itself. The high mean age at first marriage in Bali can be explained as follows. Almost all of the women who live in Bali are Hindu community. Jones (1994) reported that non Moslem societies tend to have a higher mean age of first marriage than Moslem community.

From Figure 2 we can also see that in all provinces, urban areas have higher mean age of first marriage than rural areas. It may be the consequence that in the rural society parents arrange their daughter to marry early. In contrast, in urban society women tend to spend their time for education and work before marriage. Besides, they usually do not live with their parents, so this situation causes free marriage instead of an arranged marriage (Otani 1991).

Figure 2
THE MEAN OF AGE AT THE FIRST MARRIAGE IN JAVA-BALI BY PROVINCE AND REGION

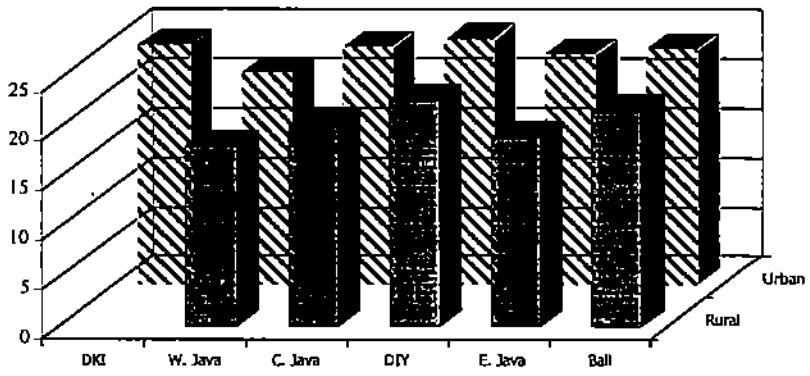
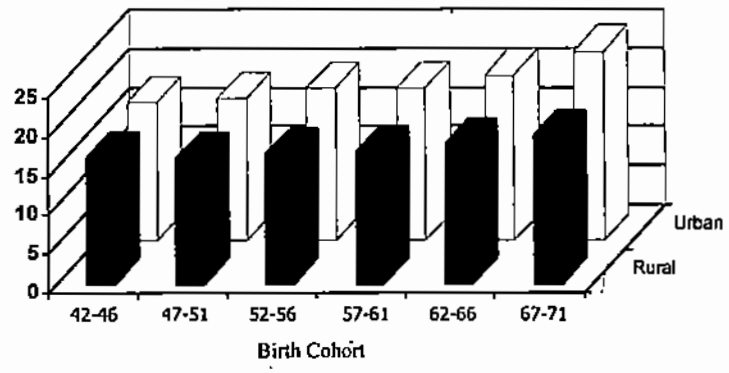


Table 2 shows the marriage pattern across cohorts. It can also be seen that the mean age of the first marriage tends to increase for younger cohort. This trend is true for both urban and rural areas (*see* Figure 3). From Table 2 can also be seen that the standard deviation is correlated with the mean age of marriage. It means that the increasing age of the first marriage does not yet happen to all member of population. Another member of the population still tends to marry early.

Figure 3
MEAN OF AGE OF THE FIRST MARRIAGE IN JAVA-BALI BY COHORT AND REGION



The incidence of marriage in Java and Bali is close to unity. Nevertheless, women lived in urban areas tend to have lower incidence of marriage for younger cohort. Conversely, women who lived in rural areas tend to have higher incidence of marriage for younger cohort. The trend of incidence of marriage for women who lived in rural area is uncommon. The lower incidences of marriage of older cohort compared to younger counterpart may be caused by economical depression during their marriage age. Another possibility is usually older women report themselves as single when they were divorced or they did not live with their spouse. Another reason is that the risk of death during delivery is quite high. Consequently, the number of women that lived until the time of interview are higher for single women than for married women. This situation causes the estimation of incidence of marriage become lower than the real patterns.

7. Conclusion

The Coale-McNeil marriage model is one of the most powerful model for measuring marriage patterns because of its ability: to adapt to conditional samples, to measure dispersion and incidence of marriage, and to smooth the data. The application of the model in Java and Bali indicated that there are a great variation of the mean age at first marriage among provinces, the incidence of marriage in Java and Bali were closed to

unity, and the standard deviation was correlated with the mean age at first marriage. The mean age at first marriage also changes with cohorts. The younger cohorts tended to marry at the older age.

Dealing with the data, we must be aware that data usually collected retrospectively. In fact, the model explains the patterns of women that survive until the time of survey. When the mortality rates for single women are different with married women, the model can not be accepted as general patterns. In the future, the availability of mortality data by marital status is needed in order to enable to examine whether the mortality and nuptiality are independent or not.

Acknowledgment

I am grateful to Drs. M. Sudarmadi of the National Family Planning and Coordinating Board, Jakarta, Indonesia, for permission to use the IDHS 1991 data.

References

Central Bureau of Statistics (CBS) Indonesia, National Family Planning Coordinating Board (NFPCB) Indonesia, Ministry of Health (MoH) Indonesia, and Demographic and Health Survey Macro International. 1992. *Indonesia Demographic and Health Survey 1991*. Jakarta.

Coale, A.J. 1971. "Age Patterns of Marriage". *Population Studies*, 25:193-214.

Coale, A.J. and McNeil, D.R. 1972. "The Distribution By Age At First Marriage In A Female Cohort". *Journal of the American Statistical Association*, 67:743-749.

Hajnal. 1953. "Age at Marriage and Proportions Marrying". *Population Studies*, 7:111-32.

Jemain, A.A. 1985. "The Maternal Mortality Pattern in Malaysian Peninsular in 1971-1980". (Corak Kematian Ibu Bersalin di Semenanjung Malaysia dari Tahun 1971-1980). *Sains Malaysiana*, 14(4):485-492.

Jones, G.W. 1994. *Marriage and Divorce in Islamic South-East Asia*. New York. Oxford University Press.

Kadi, A.S. 1987. Age at Marriage in India. *Asia-Fasific Population Journal*. 2(1): 41-56.

Otani, K. 1991. "Time Distribution in the Process to Marriage and Pregnancy in Japan". *Population Studies*, 45:473-487.

Raftery, A. E. 1986. "A Note on Bayes Factors For Loglinear Contingency Table Models with Vague Prior Information". *Journal of the Royal Statistical Society Series B*, 48(2): 249-250.

Raftery, A.E. 1986. "Choosing Models For Cross-Classifications". *American Sociological Reviews*, 51:145-146.

Rodriguez, G. and Trussell, J. 1980. "Maximum Likelihood Estimation of the Parameters of Coal's Model Nuptiality Schedule from Survey Data". *Technical Bulletins*, No. 7/TECH. 1261. World Fertility Survey.

Xie, Y. 1990. "What is Natural Fertility? The Remodeling of a Concept". *Population Index*, 56(4):656-663.

Xie, Y. and E.E. Pimentel. 1992. "Age Patterns of Marital Fertility: Revising the coale-trussell method". *Journal of the American Statistical Association* 87:977-984.

Hadi Sumarno, PhD. Lecturer at the Department of Mathematics Faculty of Mathematics and Natural Science, Bogor Agricultural University Indonesia, and staff member at the Laboratory of Applied Mathematics at the same faculty.
E-mail: ppm-mat-ipb@mipabogor.or.id; Phone: 62-251 381244; Fax: 62-251 313384.